Gfam Quantum Kernel DevKit v1.3.0 — Cancer Atlas (IC50 vs Quantum Minima)

Interactive prospectus & navigation guide — 2025

What it is. A demo web application that unifies public cell-line chemogenomic datasets (e.g., IC50 response curves) with simulated quantum minima scores to build transparent, per-sample shortlists and global leaderboards of candidate therapeutics.

What it does. For any selected cancer cohort, it ranks drugs by multiple signals — median IC50, mean quantum minima, per-sample winners — and visualizes sensitivity structure using a high-sensitivity Topographic UMAP.

How we made it. We aggregate curated drug–response tables and lineage metadata, standardize IDs, and compute two families of signals: (1) empirical wet-lab potency (IC50) and (2) a physics-inspired energy objective minimized with a stochastic optimizer (SPSA) over a kernel-parameterized landscape. The minima scores are verifiable in the sense that they are reproducible from the stated objective and optimizer traces (no hidden heuristics).

This interactive is provided for pitch/demo purposes only. It summarizes public cell-line data and experimental metrics (IC50) alongside simulated quantum minima scores (verifiable). Values are not medical advice and must not be used for patient treatment decisions. Any interpretations require independent validation in appropriate wet-lab and clinical settings. © 2025 Gfam Quantum.

Primary Sources

- DepMap / CCLE: cell-line metadata, lineages, gene expression matrices.
- Drug response tables (IC50/ActArea) from established screens (e.g., GDSC/CTRP/ALMANAC).
- Compound annotations from public resources (e.g., DrugBank / ChEMBL).
- Simulated quantum minima: kernelized objective minimized with SPSA; reproducibility artifacts (objective form, seeds, convergence traces).
- Auxiliary ontologies for disease/cancer type aliases (e.g., TCGA/NCIt term maps).

Named links (clickable)

- UMAP (dimension reduction) — McInnes, Healy, Melville — https://umap-learn.readthedocs.io/en/latest/
- UMAP preprint — "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction" — https://arxiv.org/abs/1802.03426
- DepMap Portal — downloads (includes CCLE expression/metadata) — https://depmap.org/portal/download/
- CCLE 2019 (Nature) — "Next-generation characterization of the Cancer Cell Line Encyclopedia" — https://www.nature.com/articles/s41586-019-1186-3
- GDSC / CancerRxGene — Genomics of Drug Sensitivity in Cancer — https://www.cancerrxgene.org/
- Sanger Cell Model Passports — standardized cell-line identities — https://cellmodelpassports.sanger.ac.uk/
- DrugBank — compound annotations — https://go.drugbank.com/
- ChEMBL — bioactive molecule database — https://www.ebi.ac.uk/chembl/
- NCI Thesaurus (NCIt) — controlled vocabulary — https://ncithesaurus.nci.nih.gov/ncitbrowser/
- TCGA Program — reference disease/omics resources — https://www.cancer.gov/ccg/research/genome-sequencing/tcga

Note: exact dataset versions and citations are tracked per release branch; public entry points are listed above.

Data Model & Ranking Logic (Plain-English)

Core tables.

- Cell lines with DepMap_ID, lineage/primary_disease.
- Gene expression: rows are cell lines; columns are genes (CCLE).
- Drug response: (DepMap_ID, DRUG_NAME, IC50, n).
- Quantum minima: (DepMap_ID, DRUG_NAME, quantum_minima, Q_MEAN).

Per-sample shortlists.

For each DepMap_ID, pick the row with the lowest IC50 (and separately the lowest quantum_minima). These form two shortlists per cohort, useful to compare empirical vs. simulated winners.

Global leaderboards.

- IC50 top drugs: rank by median IC50 across included samples (lower is better), tie-breaking by support count n.
- Quantum top drugs: rank by mean Q_MEAN across samples (lower is better), again tie-breaking by n.

UMAP (topographic, high-sensitivity view).

We embed gene-expression vectors to 3D UMAP, label each point by which signal wins (IC50 vs Quantum), and size/color points by a sensitivity score (percentile). High-confidence regions are highlighted using halo overlays to draw attention to stable neighborhoods.

Product Walkthrough (Screenshots)

Representative views of the application layout. Images are scaled to fit the page; captions below each figure summarize the intent.

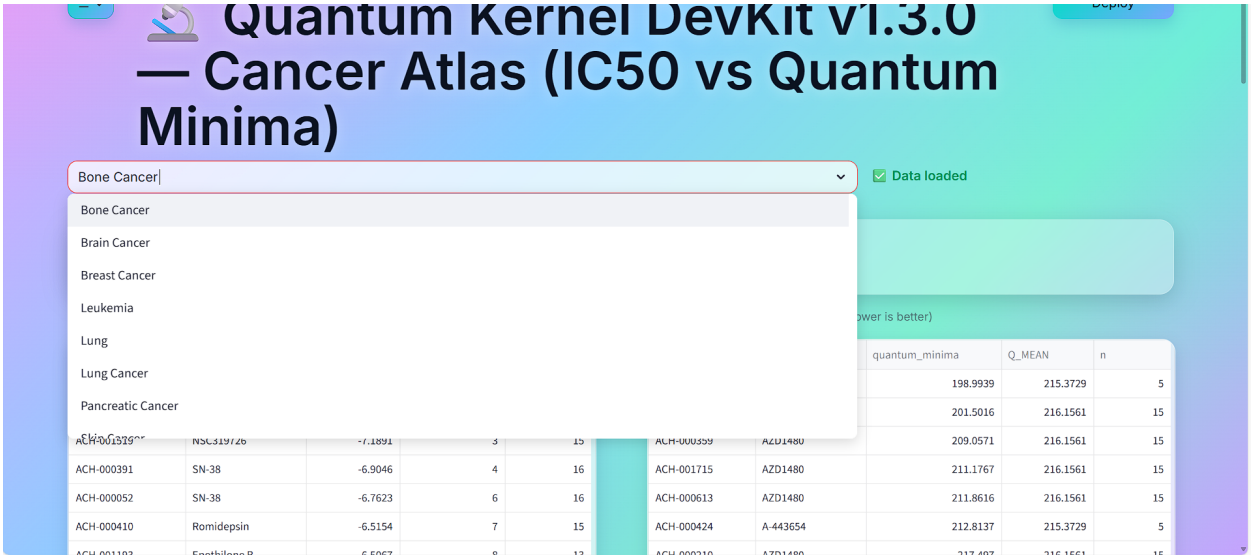


Figure 1. Application view 1.

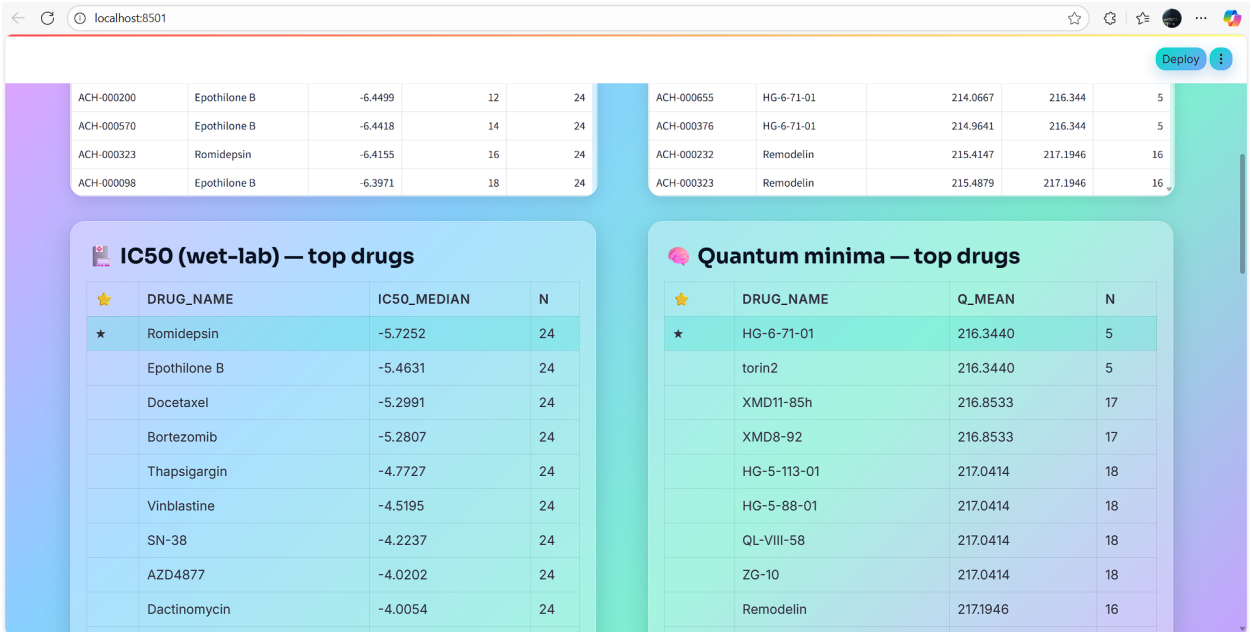


Figure 2. Application view 2.

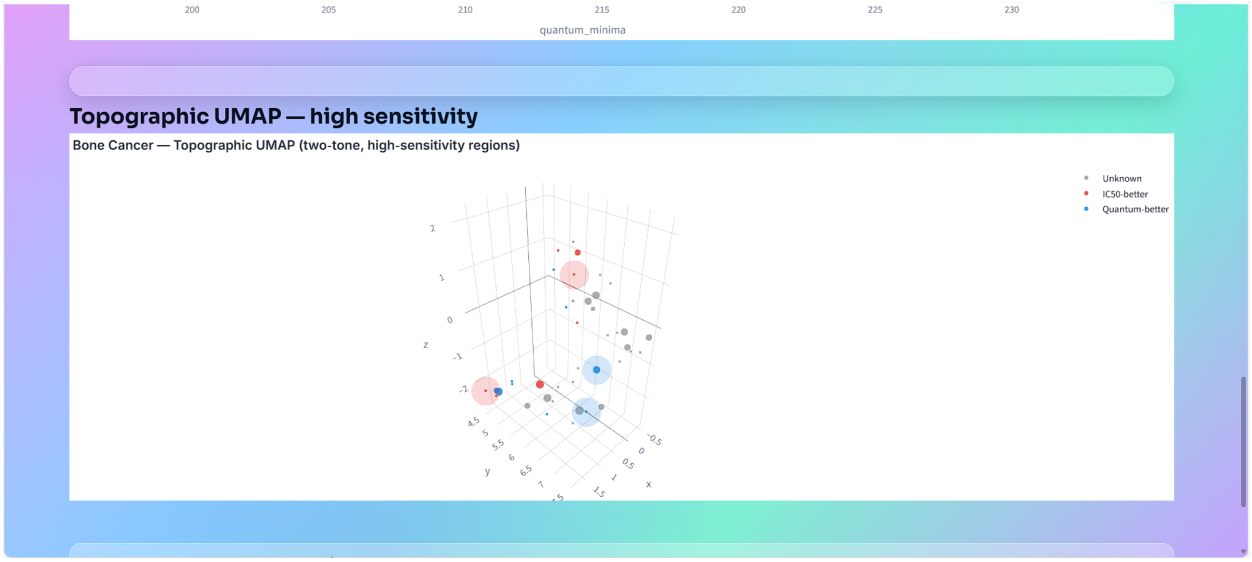
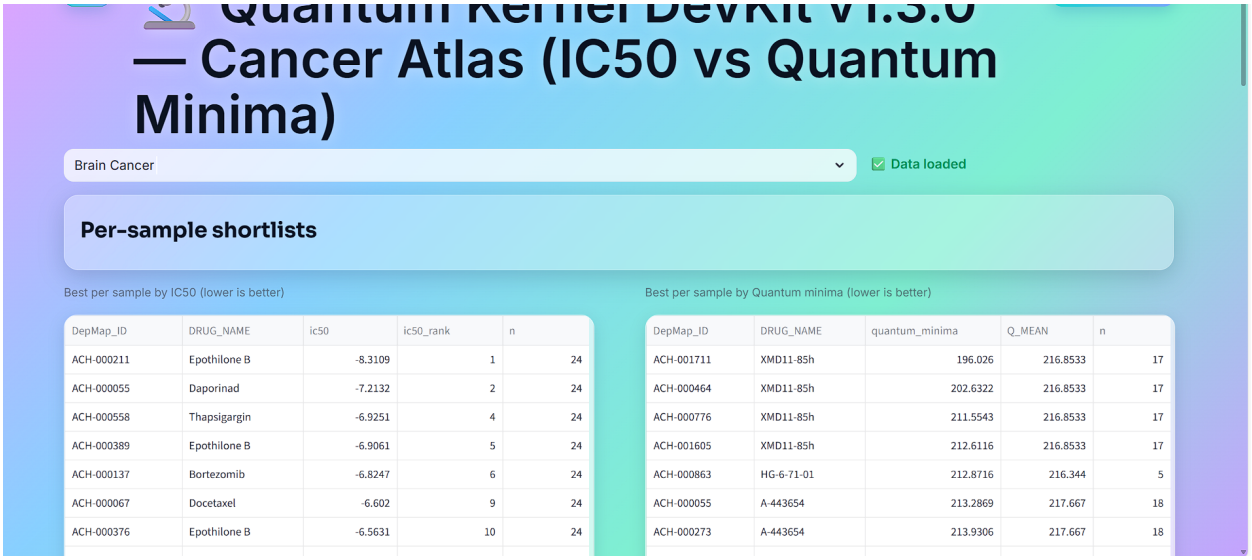


Figure 3. Application view 3.



DepMap_ID	DRUG_NAME	ic50	ic50_rank	n
ACH-000211	Epothilone B	-8.3109	1	24
ACH-000055	Daporinad	-7.2132	2	24
ACH-000558	Thapsigargin	-6.9251	4	24
ACH-000389	Epothilone B	-6.9061	5	24
ACH-000137	Bortezomib	-6.8247	6	24
ACH-000067	Docetaxel	-6.602	9	24
ACH-000376	Epothilone B	-6.5631	10	24

DepMap_ID	DRUG_NAME	quantum_minima	Q_MEAN	n
ACH-001711	XMD11-85h	196.026	216.8533	17
ACH-000464	XMD11-85h	202.6322	216.8533	17
ACH-000776	XMD11-85h	211.5543	216.8533	17
ACH-001605	XMD11-85h	212.6116	216.8533	17
ACH-000863	HG-6-71-01	212.8716	216.344	5
ACH-000055	A-443654	213.2869	217.667	18
ACH-000273	A-443654	213.9306	217.667	18

Figure 4. Application view 4.

How to Navigate the App

Select a cohort

Use the dropdown beneath the main title. Once selected, a green checkmark confirms data loaded for that cohort.

Filters

- Min n: require a minimum sample count per drug.
- Drug search: fuzzy match by prefix/substring.
- Cell search: match DepMap_ID substring.

Reading the views

- Per-sample shortlists: two tables; one selects lowest IC50 per sample; the other selects lowest quantum minima.
- Leaderboards: left = top drugs by median IC50; right = top drugs by mean quantum minima. First row is highlighted.
- Scatter: each point is a row where both signals exist; color encodes which metric wins; lower-left is generally better.
- UMAP: 3D scatter of expression; color shows winner; size reflects sensitivity percentile; halo overlay marks high-confidence regions.

Technical Appendix (Summary)

Quantum minima objective (sketch).

We define an energy-like scalar objective $E(\theta; x)$ for a sample x under a kernel-parameterized model. Minima are found by a stochastic optimizer (SPSA) directly on E. We record random seeds, step sizes, and convergence traces to permit verification/replay.

SPSA update (finite-difference, simultaneous perturbations).

At iteration k with parameters θ_k, sample a random perturbation Δ_k with entries ± 1. Evaluate E at $\theta_k \pm c_k\Delta_k$ to form a gradient estimate $\hat{g}_k \approx (E(\theta_k + c_k\Delta_k) - E(\theta_k - c_k\Delta_k)) / (2c_k) \cdot \Delta_k^{-1}$. Update $\theta_{k+1} = \theta_k - a_k \hat{g}_k$. Standard gain schedules (a_k, c_k) ensure stability.

Leaderboards & shortlists.

IC50 leaderboard: median across samples (lower is better). Quantum leaderboard: mean of minima scores (lower is better). Per-sample shortlists pick the best row within each DepMap_ID by each metric.

UMAP details.

Gene expression is optionally subsampled for performance. 3D UMAP preserves neighborhood structure; sensitivity is mapped to marker size and halos are added for the top decile.

Reproducibility.

- Fixed seeds on sampling and optimizer initializations.
- Persisted objective config and convergence traces.
- Versioned caches for cohort-level artifacts.

Note: This appendix intentionally stays high-level. Detailed kernel forms, hyperparameters, and training recipes are kept internal and can be shared under NDA.

Closing Notes

This prospectus is a guided overview of the DevKit and its rationale. For investor or partner evaluations, we can share deeper internal docs under NDA (dataset manifests, objective definitions, and validation plans).

Contact: gulfamh1999@gmail.com • LinkedIn: gulfam-hussain99 • © 2025 Gfam Quantum